Model-Based Diagnostics for Air Handling Units

Tim Salsbury and Rick Diamond
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

Introduction

In most large air-conditioned buildings, air-handling units account for a significant portion of total building energy consumption and have a major impact on comfort conditions and maintenance costs. The devices within an air-handling unit either use energy directly in the case of fans, or indirectly, in the case of heat exchangers, which impose loads on the chiller and boiler plant. Air-handling units can comprise a myriad of subsystems (fans, heat exchangers, humidifiers, mixing dampers, heat recovery units, filters, control valves, actuators, sensors, etc.) all of which are prone to developing faults. Sensor information from the various air-handling unit subsystems is typically available on a single network through an energy management and control system (EMCS). In practice, little analysis of the information on the EMCS network is carried out and the full potential of the hardware is therefore not realized. Significant potential exists for making better use of this information for data analysis and fault diagnosis in order to improve operations, save energy, and assure comfort conditions.

In order to carry out fault diagnostics, some representation (or reference) of correct or normal behavior has to be developed. This reference is the most important part of a fault diagnosis system. The consequences of a poorly defined reference are a failure to detect faults or the generation of false alarms. A model-based approach to diagnostics involves using a mathematical description of the system as a reference of correct behavior. A diagnostics scheme can use various types of models, such as first-principles models, neural networks, fuzzy rules, characteristic curves, etc. This paper advocates the use of first-principle models, which are developed from equations such as thermodynamic relationships and include parameters related to physical attributes of the real system. The paper describes how these types of models can be used as part of computer-based technologies at various stages in the life cycle of a building.

Fault Diagnostics Based on Models

Models encapsulate information about a considered system in the form of equations. The role of a model in diagnostic schemes is generally to act as a reference of expected or correct operation. Figure 1 illustrates the concept of using a model as a performance reference. In the figure, the model predicts performance of the real system based on monitored variables associated with its operation. In many diagnostic schemes, differences between model predictions and monitored system variables (known as residuals or “innovations”) form the basis of the fault diagnosis. Faults can be detected by monitoring the magnitude of the residuals and comparing them to predetermined thresholds. A more detailed analysis of the residuals and associated operating point information can be used to produce a diagnosis of the type and magnitude of a particular fault. A variant approach is to use monitored data to estimate model parameters. Faults are then detected and diagnosed by evaluating changes (residuals) in model parameters from their correct operation values.

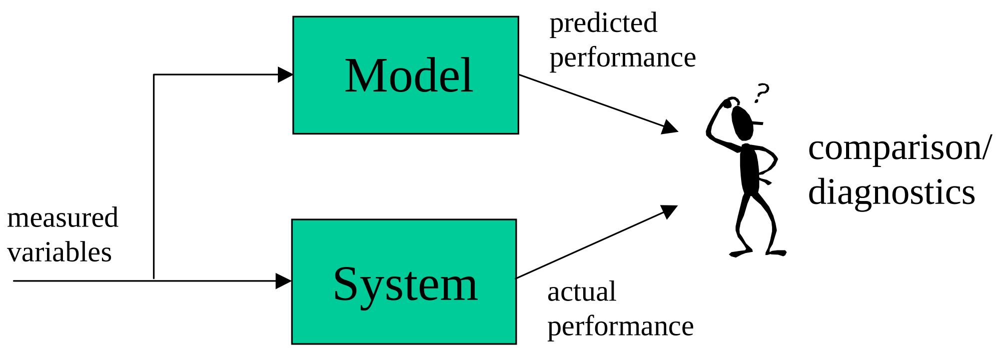

Figure 1: Concept of a model-based fault diagnostic scheme.

Several researchers (e.g., Gertler, 1998; Glass *et al.*, 1994; Isermann, 1995; Patton *et al.*, 1995) have proposed fault diagnostic schemes based on the use of models. The main trade-off with model-based schemes is configuration effort versus model accuracy. Generally, the greater the potential accuracy of the models, the greater the effort required to configure the models for use.

Advantages of using First-Principles Models

First-principles models (also referred to as analytical, white-box or physical models) are quantitative mathematical models constructed using equations derived from a theoretical analysis of the physical processes occurring in a system (e.g., heat and mass balances). This analysis determines the mathematical structure of the models and defines parameters that relate to measurable system properties. In contrast to black-box models (also known as empirical models), first-principles models make as much use as possible of prior knowledge about the system. Since the parameters of first-principles models relate to system properties, their values can, in principle, be measured directly from the real system, or obtained from information typically available in a life cycle of a building.

Supporting Infrastructures for Model-Based Applications

One of the limiting factors affecting adoption of diagnostic technology is the difficulty in configuring the methods for each individual application. Many schemes require an experienced researcher or a person with knowledge of the methods to carry out the tuning and calibration required for each real application. A large number of model-based methods, particular those based on black-box models, also require some form of training period in order to identify the reference model. These factors all add cost to the diagnostic system and limit its practical viability. The use of first-principles models alleviates some of the configuration effort since model parameters relate to real (and measurable) physical attributes. In principle, the information needed to identify the parameter values is available in the life cycle of a building.

Although the parameters of first-principles models are more easily obtainable than in other models, in most situations, limited time and resources are available to manually gather and input the parameter values. There is currently a large international effort underway in the International Alliance for Interoperability (IAI) to develop standard data models for architectural, engineering, and construction processes. Technologies based on the use of first-principles models are a potential beneficiary of such initiatives. Adoption of a standard means by which to handle data elements associated with building systems opens the way for automatic configuration of model-based technologies from other software programs used in the building life cycle. Potential sources of information include CAD, product databases, EMCS software, etc. Figure 2 illustrates the concept of a shared data model and shows how this would support model-based application software, such as diagnostics. A standard data model would act as a repository of information, which evolves during the building life cycle to reflect the actual building and its systems. Model-based applications could then utilize this repository as a basis for obtaining configuration parameters suitable for application at various life cycle processes.

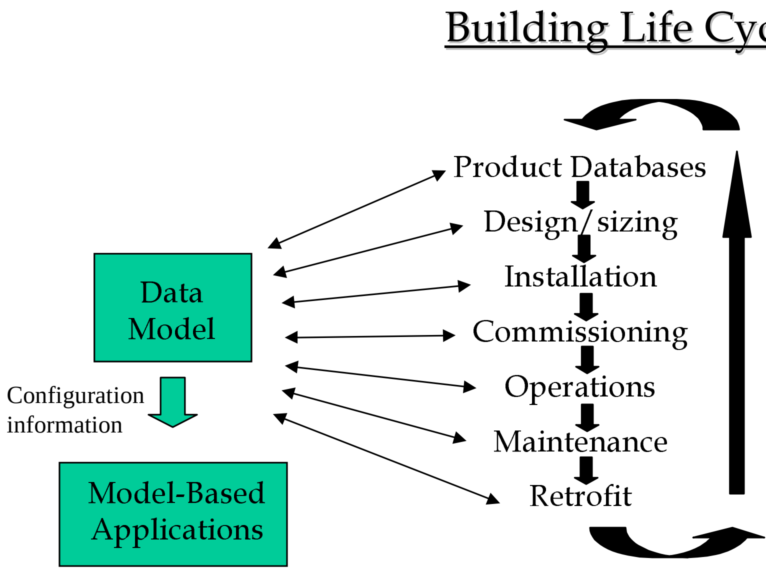

Figure 2: Concept of a standard data model in the building life cycle.

Table 1 lists examples of the kind of information required to configure a diagnostic tool developed around the use of first-principles models of an air-handling unit and its constituent subsystems. The table indicates the source of the information in terms of a life cycle process. Most of the data elements are defined at the design process, but are validated or updated at later life cycle processes.

Table 1: Example data requirements for air-handling unit models

DATA ELEMENT	UNITS	SOURCE IN LIFE CYCLE
	HEATING/COOLING COIL SUBSYSTEMS	
Maximum heat transfer rate	kW	Design
Minimum cold fluid inlet air temperature	ºC	Design
Maximum cold fluid mass flow rate	kgs^{-1}	Design/(commissioning)
Maximum hot fluid inlet temperature	ºC	Design
Maximum hot fluid mass flow rate	kgs^{-1}	Design/(commissioning)
Hot and cold fluid types	Enumeration (water, steam, etc)	Design
Flow arrangement	Enumeration (cross, parallel, counter)	Design
Valve authority	-	Commissioning
Control valve type	Enumeration Exponential, Linear)	Design/commissioning
Inlet dry-bulb temperature	ºC	Operations (EMCS)
Inlet relative humidity	%	Operations (EMCS)
Air mass flow rate	kgs^{-1}	Operations (EMCS)

FAN SUBSYSTEM (CONTROLLED TO STATIC PRESSURE SETPOINT)		
Motor efficiency at maximum load	%	Design (product data)
Outlet cross-sectional area	m^2	Design (product data)
Resistance between static pressure sensor and ambient	$kg^{-1}m^{-1}$	Design/commissioning
Inlet air temperature	ºC	Operations (EMCS)
Electric power input	kW	Operations (EMCS)
Static pressure setpoint	kPa	Operations (EMCS)/commissioning
MIXING BOX SUBSYSTEM		
Authority	-	Design/commissioning
Minimum outside air requirement	%	Design/commissioning
Return air temperature	ºC	Operations (EMCS)
Outside air temperature	ºC	Operations (EMCS)
Return air relative humidity	%	Operations (EMCS)
Outside air relative humidity	%	Operations (EMCS)

Example Air-Handling Unit

LBNL researchers have been evaluating the energy usage in the Phillip Burton federal building in San Francisco between 1996 and 1999. Part of the analysis showed the potential for energy savings from improved control and monitoring of the air-handling units. A simulation of one of the main air-handlers in the building is used here to demonstrate the potential of model-based diagnostics. The simulated system was developed in the MATLAB environment using models similar to those found in the computer simulation program HVACSIM+ (Clark, 1985).

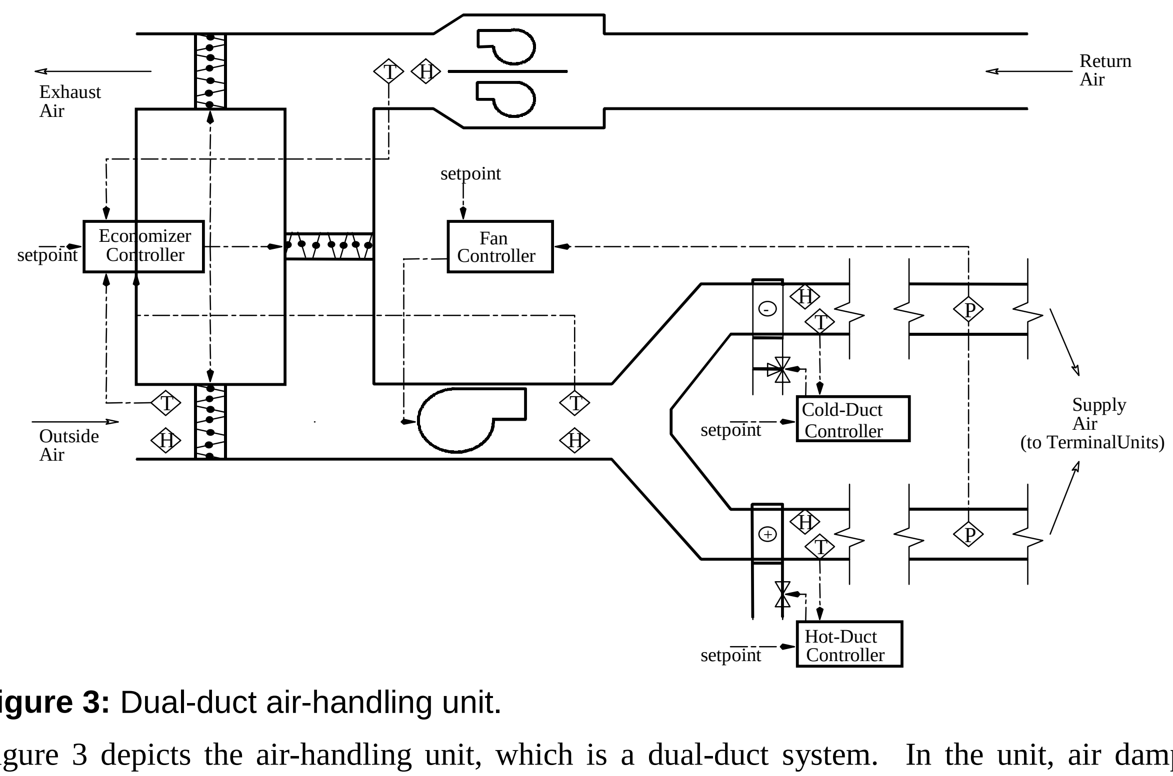

Figure 3: Dual-duct air-handling unit.

Figure 3 depicts the air-handling unit, which is a dual-duct system. In the unit, air dampers controlled by an economizer, mix return-air from the building with outside-air in order to maintain a mixed-air temperature setpoint. A large supply fan blows the mixed-air through both the hot- and cold-deck ducts. The supply fan controller maintains the average of the hot and cold ducts at a fixed static pressure setpoint. The supply fan speed varies in order to counteract changes in duct system resistance brought about by dampers opening and closing in VAV terminal units. Two fans installed in the return duct have their speeds tracked to the speed of the supply fan. The hot and cold ducts each house a heat exchanger with controllers configured to maintain setpoints by modulating control valves. The hot-duct heat exchanger has a two-port valve and the cold-duct a three-port valve. The air-handling unit has the capacity to deliver 74kg/s of air and provide 850kW of heating and 1260kW of cooling.

Use of Models in the Design Process

The design process is where much of the information needed to configure first-principles models originates. Moreover, first-principles models are the main components of simulation programs, which have now reached a level of maturity where their use in the design process of a building is a viable and increasingly widespread practice. Simulation programs are generally used for energy analysis once the final system design has been made. However, simulation is now being used more as an integral part of the design process in order to optimize the design according to given criteria such as energy use, capital cost, control performance, etc. Figure 4 illustrates the concept of using a simulation for design optimization in this way.

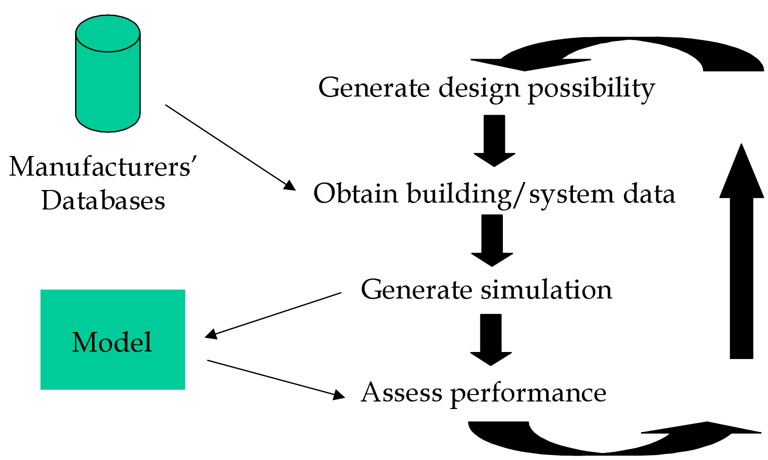

Figure 4: Simulation-assisted design concept.

Improved data handling and more powerful computing environments create the opportunity to extend the use of simulation and model-based concepts into life cycle processes beyond design. In particular, model configuration information gathered during the design process may be used in model-based technologies in life cycle processes beyond design as a way to confirm that operation conforms to design intent.

Model-Based Approach to Automatic Commissioning

The performance of many HVAC systems is limited more by poor installation, commissioning, and maintenance than by poor design (Liu, 1997; Piette, 1996; Schexnayder *et al.*, 1997). Commissioning is often carried out poorly in practice for the following reasons:

- Limited time and resources available to undertake rigorous testing

- Shortage of skilled personnel
- Difficulty in defining performance criteria for the commissioning process

An important part of the commissioning process involves carrying out a proof of operation. In large modern buildings, the energy management and control system (EMCS) is used to exercise the various systems in the building to verify electric and hydraulic connectivity, correct balancing, and proper installation. The potential exists to automate this part of the commissioning process to address the problems listed above. The benefits of an automated approach to commissioning are:

- Allows testing on systems in parallel, thereby reducing overall testing time
- Automates the labor-intensive aspects of commissioning, thereby freeing engineers to deal with problems identified by the tests
- Facilitates conformance testing and use of pre-determined test standards and performance targets

Automated commissioning involves analyzing system performance in order to detect and diagnose problems (faults) that would affect the operation of the system during normal use. LBNL researchers have developed an automated commissioning tool based on simple models. The tool is simple to configure and has the potential to detect system problems during the commissioning phase that would severely restrict performance during normal operation. Benefits are energy savings, improved occupant comfort and the avoidance of more costly maintenance during operations.

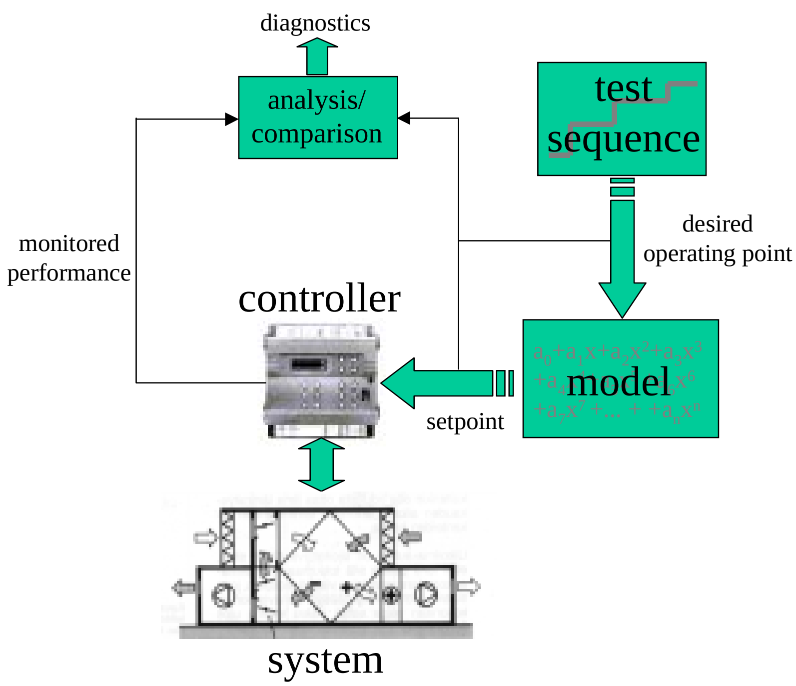

Figure 5: Model-based approach to automated commissioning.

Figure 5 illustrates the automated commissioning scheme, which tests systems while under closed-loop control. The idea is to test both the control performance and basic operation of the system simultaneously in order to reduce testing time. Operation of a particular system is tested at three operating points: high capacity, low capacity, and mid capacity. The high capacity point checks whether the system is capable of meeting design loads, the low capacity point checks for the close-off ability of valves and dampers, and the mid point checks for correct balancing. The model takes an operating point as input and makes a prediction of the setpoint expected to drive the system to that operating point. By configuring the model to represent correct operation, a deviation in the actual operating point of the system from the desired point represents an indication of faulty behavior.

Figure 6 shows results from carrying out a sequence of tests on the dual-duct air-handling unit described earlier. The top graph in the figure shows the three controlled temperatures in the air-handler: mixing plenum temperature, hot duct temperature, and cold duct temperature. The temperatures are shown as solid lines, and their setpoints as dashed lines. The lower graph shows the control signals to each of the three subsystems: mixing dampers, heating coil, and cooling coil. The test sequence for the dual-duct air-handling unit comprised six individual tests. Table 2 shows the desired operating points (control signals) for each of these tests. Note that the three operating point tests are carried out on the three different air-handling unit subsystems in only six steps by combining some of the tests. In test 1, for example, all subsystems are tested at their close-off point simultaneously.

Table 2: *Expected* control signals for the demanded setpoints. Note that indices in bold indicate figures pertinent for a particular test.

TEST NUMBER	EXPECTED CONTROL SIGNAL (%)		
	MIXING	COOLING	HEATING
1	**100**	**0**	**0**
2	100	0	**50**
3	100	0	**100**
4	**50**	0	0
5	**0**	**50**	0
6	0	**100**	0

Each change in setpoint is held for 20 minutes with the last 5 minutes of the period used to calculate two performance indices: the mean setpoint tracking error, and the average control signal. These indices are calculated over the last 5-minute period of each test since, ideally, the system is expected to be in steady state during this time. Diagnostics are generated by comparing the two indices calculated from each test with expected values. Following the test sequence, a table of results is generated of the form shown in Table 3. A diagnosis of the system performance is made by comparing the table of measured values with the expected values (Table 2) for a correctly operating system. Note that the ideal values for the mean absolute setpoint errors are all zero. Differences in the test values from their ideal values are due to inaccuracies in the models and other uncertainties in the process as a whole. In order to automate fully the diagnosis, thresholds need to be defined for the test indices. Since the types of problems at commissioning time typically lead to large disruptions in performance (e.g., reverse acting actuators, disconnected valves, etc.), reasonable estimates for the thresholds can be made from expert knowledge. This approach has been investigated in related work (Haves et al., 1996) where fuzzy inferencing was used to assign ranges of confidences for index values.

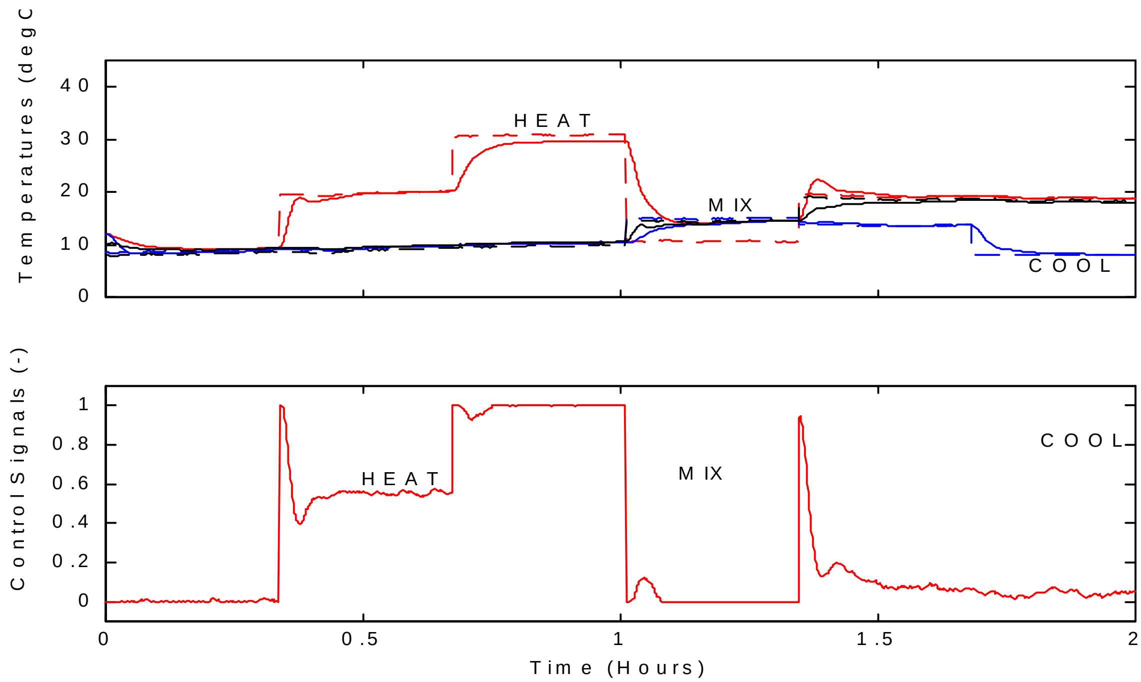

Figure 6: Results from commissioning tests on a correctly operating system.

Table 3: Results of tests on correctly operating system. Note that indices in bold indicate figures pertinent for a particular test.

Test Number	Average Control Signal (%)			Mean Absolute Error (K)		
	Mixing	Cooling	Heating	Mixing	Cooling	Heating
1	**99**	**7**	**0**	**0.7**	**0.2**	**0.3**
2	99	2	**55**	0.6	0.1	**0.1**
3	98	2	**100**	0.5	0.2	**1.3**
4	**52**	0	0	**0.1**	0.7	4.0
5	**2**	**38**	7	**0.4**	**0.0**	0.1
6	2	**92**	5	0.3	**0.0**	0.1

Model-Based Diagnostics during Operations

Heating, ventilating, and air-conditioning (HVAC) systems are typically controlled using proportional plus integral (and sometimes plus derivative) PI(D) control law. In practice, HVAC systems exhibit non-linear operating characteristics. Because PI(D) control law is best-suited to linear systems, the performance of an HVAC controller can vary when operating conditions change. A controller tuned for one set of environmental conditions may become sluggish or oscillatory at a different set of conditions (Salsbury, 1998). Poor control performance can lead to occupant discomfort in a building, greater energy consumption, and increased wear on controlled elements, such as actuators, valves, and dampers.

In a conventional PI(D) feedback loop, the controller does not contain much information about the process it is controlling. Faults that lead to performance deterioration or a change in system behavior are often masked by the feedback loop. One way to address the problem of inconsistent control performance and be able to detect faults in the controlled process is to use a model of the correctly operating system as part of the control scheme. The model is combined with the feedback loop to reduce the effects of plant non-linearity on control performance, while the feedback loop compensates for modeling errors and improves the response time. If the system under control deviates from correct operation, the feedback controller compensates and effectively provides a measure of the residual between model and system performance.

Figure 7 shows a schematic diagram of the combined fault detection and control scheme. In the scheme, the control signal issued by a conventional HVAC controller (u_{PI}) is supplemented by a control signal generated by the model (u_{FF}). The model is in inverse form and produces a control action appropriate for the current setpoint and measured disturbances. Since the model is configured to represent the correctly operating system, the output from the conventional (in this case PI) controller (u_{PI}) is a measure of the residual between predicted and actual performance. This residual forms the basis of fault detection whereby the scheme generates an alarm when the magnitude of the residual exceeds a threshold.

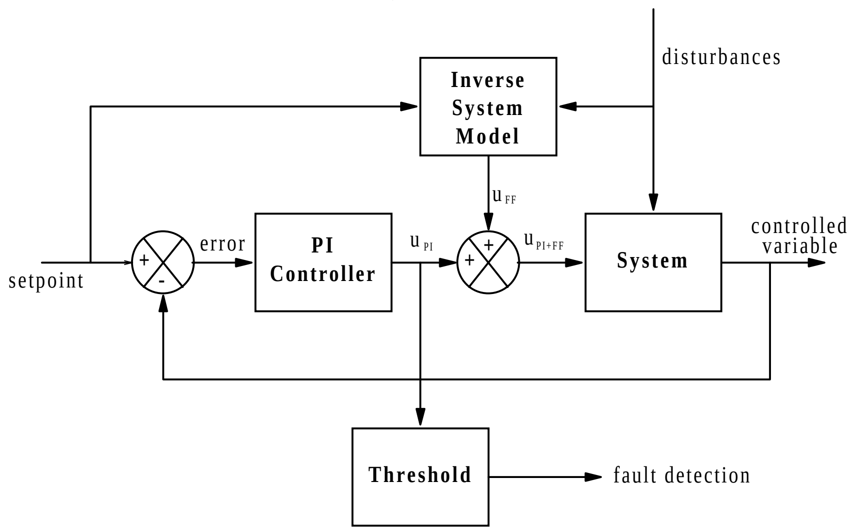

Figure 7: The control and fault detection scheme.

The potential of using the control scheme for fault detection is illustrated via example results obtained from tests on a simulation of the dual-duct system described earlier. The scheme was used to detect a stuck damper in the mixing box, which is a common fault in HVAC systems. This fault can be due to various causes, such as a failed actuator, damper obstruction, de-coupled linkage, etc. The fault is difficult to detect in practice as it does not cause failure of the mixing process but instead alters its characteristics and restricts the operating range.

Twelve hours of test data, sampled at 5-second intervals are used to test the control scheme. The test data contain real measurements of ambient and return air temperatures from a real building. Airflow rate is artificially varied in a sequence of arbitrary steps and ramps. The setpoints for each of the three subsystems are also varied artificially in order to facilitate a rigorous assessment of the controllers. Figure 8 shows the control performance when the re-circulation damper is stuck at 50% open. The upper graph shows the three controlled variables and setpoints. The middle graph shows the ambient and return air temperatures and the air flow rate (right axis scale), and the lower graph shows the control signals to each of the subsystems. The fault is masked for the majority of the test data with the fault condition only affecting control performance in the middle of the data when the controller demands 100% outside air.

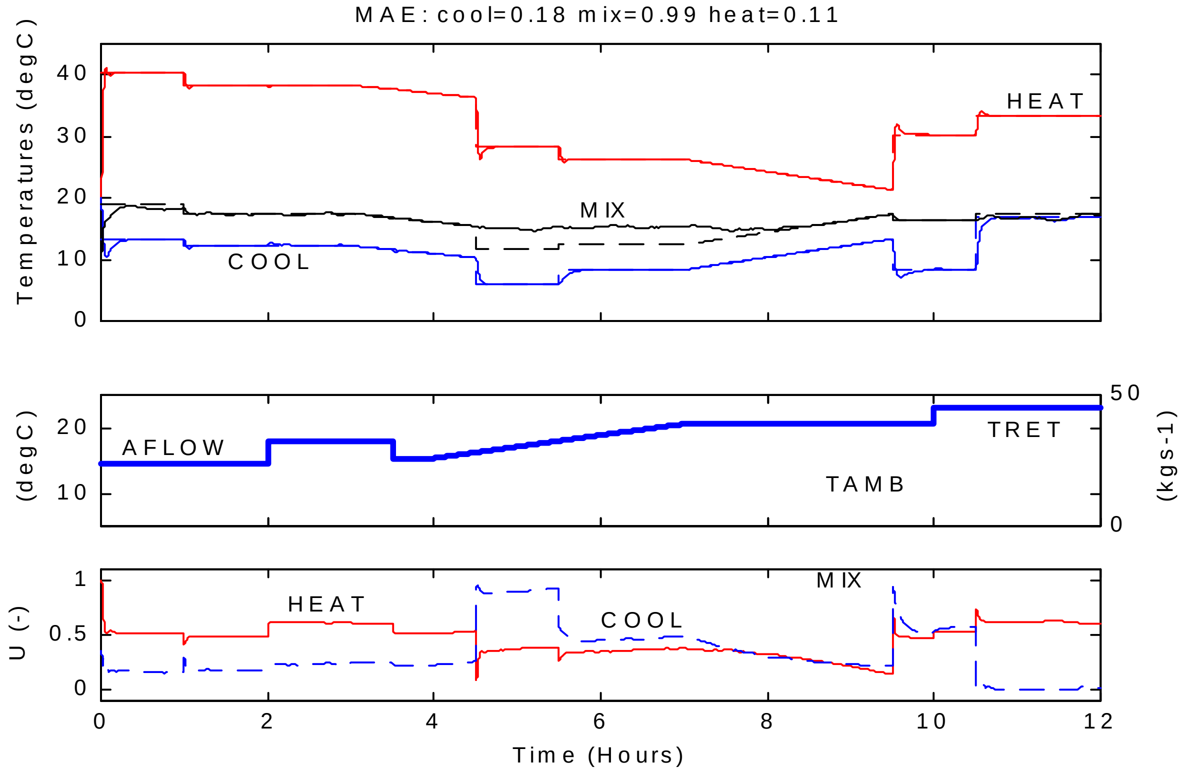

Figure 8: Control behavior when the return air damper is stuck at 50%.

Figure 9 shows the fault indicator variables for the stuck damper condition. The top graph shows the control signals generated by the three feedback loops in the air-handler (i.e., the residuals between model predictions and the observed behavior). If the models used in the controller were perfect representations of the system under control, these control signals would asymptotically approach zero following each disturbance. Figure 9 shows that there are modeling errors in each of the controller models. The errors vary during the data due to transient effects and because the models approximate the system better in certain parts of the operating range than in others. A fault detection threshold was determined based on model accuracy by testing the controller with the correctly operating system at various operating points. The feedback control action to the mixing box increases during the periods when the dampers are supposed to be providing 100% outside air. Unwanted re-circulation through the mixing box reduces its operating range so that setpoints close to the ambient air temperature become unattainable. The fault is less apparent in the parts of the range where setpoints are closer to the return air temperatures. The fault proves difficult to detect and the threshold is only exceeded for a short time around nine hours into the test data. However, the period over which the feedback action remains above the threshold is sufficient to trigger an alarm for the mixing box. Note that 30 minutes is set as the time limit for a threshold transgression in order to trigger an alarm.

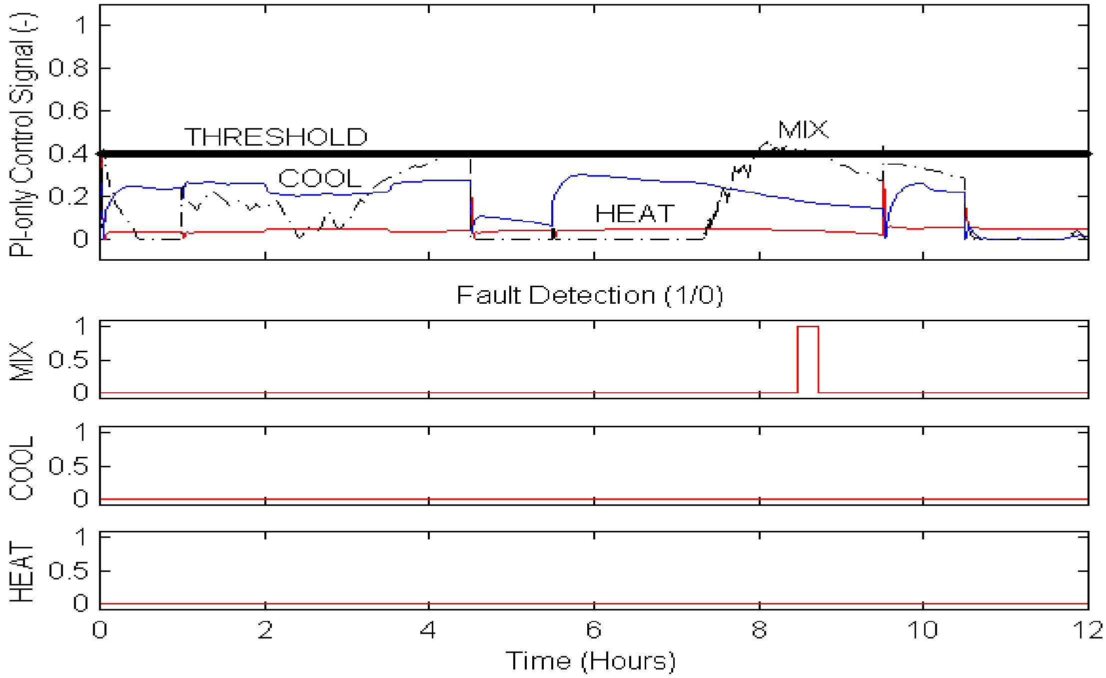

Figure 9: Fault indicator variables - return air damper stuck at 50%.

The concept of using a model for the combined purpose of control and fault detection is a logical integration of technologies as control problems are prevalent in HVAC systems and frequently lead to failures of controlled elements. In addition, since the described approach represents an "add-on" to current control technology (such as PI control), implementation in existing devices is viable. LBNL researchers have also developed diagnostic schemes for operations based on models acting in isolation to the control (Salsbury *et al*., 1995; Haves *et al*., 1996). In these approaches, fault detection is performed by non-intrusively comparing model predictions with monitored data from the considered system.

Conclusions

This paper has described how first-principles models can be used for three applications in the life cycle of a building: simulation-assisted design, automated commissioning, and operational diagnostics. Results were presented from tests carried out on an example air-handling unit to illustrate the potential of model-based technologies at the commissioning and operations stages.

Maintaining models as the central theme of applications across the life cycle of a building reduces the potential for information loss by allowing performance information to perpetuate and evolve through life cycle processes. Standardized data models facilitate easier data exchange between software programs and increase the practicability of model-based technology.

Acknowledgements

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Building Technology and Community Systems, and the Federal Energy Management Program, of the US Department of Energy under Contract No. DE-AC03-

76SF00098. The authors would like to thank Philip Haves, Robert Hitchcock, and Mary Ann Piette for their contributions to this paper.

References

Clark, R. C. 1985. “HVACSIM+ Building Systems and Equipment Simulation Program Reference Manual”. Published by the U.S. Department of Commerce, National Bureau of Standards, National Engineering Laboratory, Center for Building Technology, Building Equipment Division, Gaithersburg, MD 20899.

Gertler, J. 1988. “Survey of Model-based Failure Detection and Isolation in Complex Plants”. IEEE Control Systems Magazine. Number 6. Volume 8. Page 3.

Glass, A. S., P. Gruber, M. Roos, and J. Todtli. 1994. “Preliminary Evaluation of a Qualitative Model-Based Fault Detector for a Central Air-Handling Unit”. Proceedings of 3rd IEEE Conference on Control Applications, Glasgow.

Haves, P., D. R. JØrgensen, T. I. Salsbury, A. L. Dexter. 1996. “Development and Testing of a Prototype Tool for HVAC Control System Commissioning”. *ASHRAE Transactions*. Number 102. Part 1.

Haves, P., T. I. Salsbury, J. A. Wright. 1996. “Condition Monitoring of HVAC Subsystems using First-Principles Models”. *ASHRAE Transactions*. Number 102. Part-1.

IAI – International Alliance for Interoperability web site: *iaiweb.lbl.gov*.

Isermann, R. 1995. “Model-based Fault Detection and Diagnosis Methods”. Proceedings of the American Control Conference, Seattle, Washington, USA. Page 1605.

Liu, M., Claridge, D., Haberl, J., and D. Turner, "Improving Building Energy System Performance by Continuous Commissioning," Proceedings of the Fifth National Conference on Building Commissioning, April 28-30, 1997, Long Beach, CA.

Patton, R. J., P. Frank, R. Clark. 1989. “Fault Diagnosis in Dynamic Systems: Theory and Application”. Published by Prentice Hall.

Piette, M.A. and B. Nordman, 1996. "Costs and Benefits of Utility Funded Commissioning of Energy-Efficiency Measures in 16 Buildings," ASHRAE Transactions, Atlanta, GA, Vol. 102, Pt 1. Feb. 1996, LBNL-37823.

Salsbury, T. I. 1998. “A Temperature Controller for VAV Air-Handling Units Based on Simplified Physical Models”. ASHRAE HVAC&R Research Journal. Volume 4. Number 3.

Salsbury, T. I., P. Haves, J. A. Wright. 1995. “A Fault Detection and Diagnosis Method Based on First-Principles Models and Expert Rules”. Proceedings of Tsinghua HVAC-95. Beijing, Peoples Republic of China.

Schexnayder, A., Lunneber, T., and Ring, E., "Energy Resource Center Commissioning Results", Proceedings of the Fifth National Conference on Building Commissioning, April 28-30, 1997, Long Beach, CA.